

GEO CITATION RESEARCH

WHAT AI MODELS ACTUALLY CITE.

We analyzed 21,571 citations from Google AI Overview, ChatGPT, Perplexity, Claude, Google AI Mode and Gemini, plus a parallel experiment of 518 searches. We found eight findings, and five points where our data does not match what the industry keeps repeating.

21,571

URL CITATIONS ANALYZED

14,414

UNIQUE URLS

7,282

DISTINCT DOMAINS

6

AI MODELS ANALYZED

EIGHT FINDINGS THAT MATTER MORE THAN THEY SEEM

Most published GEO studies count mentions. We count mentions, positions and surfaces at the same time, and that is where the picture changes.

BEFORE WE START • TWO DEFINITIONS

A **citation** is every link to a web page that an AI model shows as a source for its answer. If an answer lists seven sources, we count seven citations. It is the unit of this whole study.

A **surface** is the place where that answer appears. Google has two that behave very differently: the AI Overview, the summary shown above the results the moment you run a search, and AI Mode, the follow-up chat you move into when you ask again. A good part of this document exists because those two surfaces were being measured as if they were one.

FINDING 01

The typical citation is an internal page, not a site's homepage.

Between 67.8% and 75.2% of citations point to pages two or more levels deep inside the site. 99.98% are HTTPS, 94.1% do not end in a file extension and only 4.2% carry parameters. Models cite clean, specific pages, not homepages.

FINDING 02

The blog wins volume and loses position. It is the metric nobody publishes.

The corporate blog is the most cited page type by count, with 8,139 citations with a known position, but only 7.5% reach the first place in the source list, and its average position is 10.2. Reference sites, with twenty times less volume, reach first place 24.7% of the time.

FINDING 03

Four models eat from the same plate. ChatGPT eats from another.

The main course for Claude, Gemini, Perplexity and Google AI Mode is the company blog, between 33.1% and 42.5% of their citations. ChatGPT's is the product page, at 31.3%, and its blog barely reaches 12.6%. Any pair of models we take agrees on 3.5% of the exact URLs it cites.

FINDING 04

The conversation has no owner, and no brand site controls it.

The most cited domain in the whole study accounts for 3.6% of the volume, and it takes about 500 domains to reach half. Within a specific category, no brand site exceeded 10% of the citations in its own field: 96.7% go to third parties.

FINDING 05

The AI Overview repeats your SEO. AI Mode discards it.

In 518 searches where we captured the four surfaces of the same results screen, 47.45% of the AI Overview citations were already in the organic results of that same search. When you ask again and move into AI Mode, only 19.98% of those organic pages survive.

FINDING 06

Paid ads do not exist for the generative answer.

Ads were 0.03% of AI Overview citations. And persistence into AI Mode was 0%, without a single exception across ten industries. Search ad spend is not a lever for AI visibility.

FINDING 07

The AI Overview overweights video far beyond SEO.

YouTube weighs 1.72% in organic results and 7.72% in AI Overview citations: more than four times. Added to social networks it reaches 16.08%. Reddit and Wikipedia together, in the same sample, add up to 1.57%.

FINDING 08

AEO has a real floor and no ceiling.

Of the pages with an F score, meaning below 40 out of 100, none was ever cited. Above that threshold we see no difference: pages with a C grade are cited seven times more often than pages with a B. There is no perfect number to chase, there is a floor to cross.

01 • METHOD

THREE DATASETS WE DO NOT MIX

DATASET	WHAT IT MEASURES	CITATIONS	URLS	DOMAINS
Multi-model	Monitoring queries across the five models	10,618	6,007	3,355
Live Google	Real AI Overview and AI Mode captures	10,953	8,714	4,452
Parallel experiment	518 searches with the four surfaces together	10,913	8,247	n/a

The first two datasets add up to the 21,571 citations in the title. The third, which we analyze in the Google surfaces section, was captured with the same instrument and in a window almost identical to the second, so it is not an independent replication and is not added to the total: its 8,247 URLs are the ones cited in the AI Overview (3,899) and the ones present in the organic results (4,348) of those same searches. The overlap between the first two datasets is only 307 exact URLs, so they are different experiments and are reported separately.

WHAT COUNTS AS A CITATION

Every time a model shows a link as a source for its answer, that is a citation. If the same page appears in two different answers, that is two citations. If an answer lists seven sources, that is seven citations, and we also record where in that list each one appeared. That data point appears in almost no published study and it is the one that supports the finding about the blog.

CORPUS NOTE

Gemini citations that point to internal Vertex redirects are excluded from the whole analysis, because they are temporary links that cannot be resolved after capture. That is 1,914 citations. 10,618 usable citations remain in the multi-model dataset. The queries in the first two datasets come from monitoring programs, meaning they measure how models respond to a controlled stimulus, not a user's spontaneous demand.

02 · THE SHAPE OF A CITATION

CLEAN INTERNAL PAGES, ALMOST WITHOUT EXCEPTION

WHAT URL DEPTH MEANS

It is how many levels you have to go down from the homepage to reach the page.

example.com → depth 0 · the homepage

example.com/blog → depth 1 · a section

example.com/blog/how-to-measure-geo → depth 2 · a specific page

This matters because it separates two very different things: a model citing your company (the homepage) or citing a specific answer you published (an internal page). The second is far more frequent, and that changes where it pays to invest.

SIGNAL IN THE CITED URL	VALUE	DEPTH DISTRIBUTION		
Uses HTTPS	99.98%	Usable citations · multi-model dataset		
No file extension	94.1%	0 · homepage		847
Depth 2 or more (multi-model)	67.8%	1 · section		2,574
Depth 2 or more (Google)	75.2%	2 · page		4,300
With parameters in the address	4.2%	3		1,680
With an extension (.pdf, .html...)	5.9%	4		704
		5 or more		513

The peak at depth 2 is sharp and repeats across two independent datasets that share only 3.7% of their URLs. That gives it robustness: it is not an artifact of one instrument or one category.

EXTERNAL LITERATURE

CONVERGES, FROM ANOTHER ANGLE

The University of Toronto study presented at EDBT/ICDT 2026 documents that generative answers and search results diverge significantly in the domains consulted, in the type of those domains and in the freshness of the information.²

Ahrefs reports that 28.3% of the pages most cited by ChatGPT have no organic visibility in Google at all.¹¹

OUR READING

THE IMPLICATION IS BORING AND HARD TO SELL

If 94.1% of citations do not end in an extension and two out of three live at depth 2 or more, preparing for AI is largely architecture hygiene: clean, stable, specific addresses.

There is no trick here. There is an entry requirement that many sites still do not meet.

EVERYONE MEASURES MENTIONS. ALMOST NOBODY MEASURES POSITION.

This is the finding that most separates us from the published literature, and not because it contradicts its numbers, but because it adds a dimension those numbers do not have.

WHAT THE POSITION OF A CITATION IS

When a model answers, it lists its sources in an order. That order is neither random nor cosmetic: the first source is the most visible and, in many interfaces, the only one you see without expanding the rest. When we say a page type reaches first place 24.7% of the time, we mean that out of all its appearances, in 24.7% it was the number one source of the answer.

A content type can dominate the total citation count and, at the same time, be systematically pushed to the bottom of the list. That is exactly what happens with the blog.

PAGE TYPE	CITATIONS	% IN 1ST PLACE	% IN TOP 3	AVG. POSITION
Homepage	1,165	14.2%	40.9%	5.9
Guide or resource	2,436	12.5%	37.4%	8.4
Documentation	265	12.1%	39.6%	8.1
Marketplace listing	165	12.1%	47.3%	6.2
Social and video	1,835	11.4%	37.1%	7.5
Product page	4,841	8.6%	30.7%	8.5
News article	1,332	8.0%	27.0%	9.9
Company blog	8,139	7.5%	26.0%	10.2
PDF	411	6.8%	33.1%	7.4
Comparison	402	5.2%	21.9%	10.0

Average position is the mean place that page type occupies within the source list of an answer. The lower, the better. Categories with at least 100 citations with a known position are included.

THE BLOG IS THE TICKET IN. IT IS NOT THE FRONT SEAT.

WHERE WE AGREE WITH THE LITERATURE

On volume, we agree. BuzzStream's analysis holds that blog-style editorial content accounts for more than half of all AI citations,¹⁹ and ZipTie reports that Claude favors blogs in 43.8% of its main citations.⁸ Our data shows Claude as the model most inclined to the blog, at 42.5%, and that page type is indeed the most cited in both datasets.

WHERE WE DO NOT

None of those studies publishes position. And once you add it, the blog drops to the bottom of the table: 7.5% first places against 14.2% for homepages and 24.7% for reference sites, with an average position of 10.2.

This matters because order is almost all there is. Pew Research found that only 1% of users click any link inside an AI Overview.²⁴ If the click is that scarce, the value of the citation concentrates in the visibility of the place it occupies.

SITE TYPE SORTS DIFFERENTLY THAN PAGE TYPE

The previous table looks at what the page is. This one looks at what kind of site it comes from, which is a different question and produces a different order. Institutional and reference sources have little volume and a lot of positional authority. Commercial ones, the opposite.

SITE TYPE	CITATIONS	% IN 1ST PLACE	% IN TOP 3	AVG. POSITION
Reference (encyclopedia - type)	396	24.7%	58.1%	4.7
Government	616	17.2%	49.5%	5.5
Video platform	789	13.9%	43.5%	6.7
Marketplace	267	12.4%	48.7%	5.9
Social network	827	10.9%	35.2%	7.4
Media and news	544	9.2%	29.8%	9.5
Company site	17,536	8.4%	28.9%	9.3
Academic	265	6.0%	30.9%	7.7
Professional network	219	4.6%	21.5%	11.0

Company sites are 81% of total volume and eighth of nine in first-place rate. Reference sites are 1.8% of the volume and first in the table. That contrast is the same one that orders the previous section, seen from the other axis.

DASHCRAB HYPOTHESIS 01

"The blog dominates AI citations" and "the blog wins AI citations" are different claims, but the industry uses them as synonyms.

We propose separating two indicators that are fused today: reach (do we appear?) and quality of place (do we appear first?). The blog optimizes the first and is mediocre at the second. Reference, government and marketplace pages optimize the second with a fraction of the volume. A content strategy that only measures mentions will overinvest in blog for years without noticing it is buying the wrong seat.

This does not say stop blogging. It says: stop measuring mentions alone, and do not expect the blog to give you first place, because structurally it does not.

A CONTRAST WORTH LOOKING AT TWICE

The company blog has **twenty times** the volume of reference sites and less than a third of their first-place rate: 8,139 citations against 396, and 7.5% against 24.7%. The question this opens is not whether it pays to write. It is what kind of asset you are building when you write.

FOUR MODELS EAT FROM THE SAME PLATE. CHATGPT DOES NOT.

We classified every citation by the type of page it points to and by the type of site it comes from. 97% was identified. That is where dishes the coarse classification could not see show up.

MODEL	MAIN COURSE	WEIGHT	ITS BLOG	CITATIONS	% DEEP	AVG. POSITION
Claude	Company blog	42.5%	42.5%	2,909	66.7%	5.4
Google AI Mode	Company blog	39.1%	39.1%	2,514	69.8%	10.8
Gemini	Company blog	38.7%	38.7%	1,567	58.5%	12.4
Perplexity	Company blog	33.1%	33.1%	1,477	70.4%	7.3
ChatGPT	Product page	31.3%	12.6%	2,151	71.9%	3.6

PAGE MIX BY MODEL, AS A PERCENTAGE OF ITS CITATIONS

MODEL	BLOG	PRODUCT	GUIDE	HOME	SOCIAL AND VIDEO	NEWS ARTICLE	PDF	DOCS	PRICING
ChatGPT	12.6	31.3	15.4	10.6	0.7	5.7	9.1	3.9	2.5
Claude	42.5	28.2	8.9	6.9	0.5	5.2	0.4	0.8	0.1
Gemini	38.7	22.4	7.6	16.5	3.1	4.9	0.2	0.5	0.6
Perplexity	33.1	30.6	10.6	7.0	4.7	2.8	2.6	0.9	0.7
Google AI Mode	39.1	23.4	6.9	5.6	12.1	6.1	0.5	0.8	0.2

DOMAIN MIX BY MODEL, AS A PERCENTAGE OF ITS CITATIONS

MODEL	COMPANY	GOVERNMENT	MEDIA	REFERENCE	MARKETPLACE	SOCIAL NETWORK	VIDEO	LINKEDIN	ACADEMIC
ChatGPT	76.4	9.6	3.2	4.0	3.5	0.6	0.0	0.1	1.9
Claude	92.3	0.8	1.5	2.2	1.2	0.2	0.2	0.1	1.0
Gemini	89.9	0.7	1.1	1.5	1.7	1.5	1.6	0.1	0.8
Perplexity	86.5	3.8	1.4	0.9	0.7	1.6	1.7	1.5	1.1
Google AI Mode	82.3	1.6	1.2	0.7	1.1	6.8	3.6	1.7	0.7

FIVE PORTRAITS

CHATGPT - 2,151 CITATIONS

The odd one out. It is the only one whose main course is not the blog: product 31.3% against blog 12.6%, the lowest of the five. Adding product, guides, PDFs, documentation and pricing it reaches 62% of its citations. It is also the one that cites the most PDFs, at 9.1% against 0.2% for Gemini, the one that uses the most government sources, at 9.6%, and the one that places its sources best, at an average position of 3.6. **It is the only model you serve better with documentation than with the blog.**

CLAUDE - 2,909 CITATIONS

The most editorial and almost a mono-diet: 92.3% of its citations go to company sites, the highest share in the group, and the corporate blog takes 42.5%. It practically never touches government (0.8%), social (0.2%) or PDFs (0.4%). This is consistent with what ZipTie reports, which measures 43.8% blogs in its main citations.⁸ Here the article is the unit and everything else is noise.

GEMINI - 1,567 CITATIONS

The one that cites the homepage the most, at 16.5%, and the least deep in the group: 58.5% internal pages against 71.9% for ChatGPT. Also the one that places its sources worst, at an average position of 12.4. It is the only model where a site's homepage is a real citation asset and not an accident.

PERPLEXITY - 1,477 CITATIONS

The most balanced profile: blog 33.1% and product 30.6%, almost tied, and some of everything else. It is the second-biggest citer of LinkedIn, at 1.5%, and second in government sources, at 3.8%. It is also half of the most overlapping pair in the study.

GOOGLE AI MODE - 2,514 CITATIONS

The only one that seriously eats platforms, with the breakdown in plain sight: of its 12.1%, 6.8% is social networks, 3.6% video and 1.7% LinkedIn. Claude, in the same three categories, adds up to 0.5%. That is a twenty-four-fold difference. Its most cited domains are YouTube, Instagram, Reddit and LinkedIn, in that order.

The 3% of citations that fit no named category remains unclassified. The previous classification in this study, which we inferred from the page address, left more than half in that bag.

THERE IS NO URL FOR EVERY LLM

HOW OVERLAP IS MEASURED

The Jaccard index is a simple way to compare two lists: you divide what they have in common by the total number of distinct elements across both. If two models cite exactly the same pages, it gives 1. If they share none, it gives 0.

Our median value is **0.035** for exact URLs, meaning any pair of models we take agrees on barely 3 or 4 out of every 100 cited pages, and **0.104** at the domain level. In other words: they may trust the same sites, but they rarely pick the same page from those sites.

OVERLAP ACROSS THE TEN POSSIBLE PAIRS

No pair exceeds 0.16 out of a possible maximum of 1.00. Bars at real scale.



Domain-level values. At the exact-page level the same calculation runs from 0.013 to 0.093, with a median of 0.035. Google's two surfaces, Gemini and AI Mode, share 0.094: less than Claude with Perplexity, which belong to different companies.

CONVERGENCE WITH THE LITERATURE • HIGH

This is the most replicated finding in the field, and the most ignored by budgets.

Writesonic, over 161,286 prompts, finds that models share around 17% of their sources and that only 3.8% is universal; its most overlapping pair does not exceed 0.25.⁵ Profound, over 100,000 queries, reports close to 89% distinct citations between ChatGPT and Perplexity.⁶ Averi, over 680 million citations, measures 11% domain overlap between those two models.⁷ The work by Chen and colleagues documents that services differ from each other in domain diversity, freshness, stability across languages and sensitivity to phrasing.¹

Our values fall within that range and push it toward the low end. The likely difference: we measure exact URL as well as domain.

THE COMMERCIAL CONSEQUENCE IS DIRECT AND UNCOMFORTABLE TO SELL: YOU CANNOT PROMISE TO OPTIMIZE ONE PAGE FOR EVERY LLM.

What you can build is a domain with a portfolio of pages, and measure model by model. And with the taxonomy in hand, that portfolio is no longer a vague idea: for Claude it is the blog, for ChatGPT it is product documentation, for AI Mode it is video and social.

NOBODY OWNS THE CONVERSATION. AND THAT CHANGES HOW TO PLAY.

The citation market has no dominant winner. It has thousands of small participants, and that has direct consequences for what you can expect from your own site.

CUMULATIVE DOMAIN CONCENTRATION

Usable citations from both datasets • 7,282 distinct domains competing



Not even YouTube, the most cited domain in the study, reaches 4% of total volume. It takes about 500 domains to cover half. There is no oligopoly: the long tail is the market.

TWO PERCENTAGES THAT LOOK CONTRADICTIONARY AND ARE NOT

The **3.6%** is the share of the most cited domain over all the citations in the study, mixing every category and every question. The denominator is 21,571 citations spread across 7,282 domains.

The **9.5%** that appears further down is something else: the maximum share a brand site reached within the questions of its own category, where a few dozen relevant domains compete. It is a much smaller and much more favorable denominator.

That such a low ceiling shows up even in the favorable scenario is precisely what makes the data point strong.

THE CEILING OF YOUR OWN SITE

When you measure how many of a category's citations point to the site of one of that category's brands, the number is systematically low. Across every category we analyzed, **no brand domain exceeded 10% of the citations in its own field**, with an observed maximum of 9.5% and an average of 3.3%. The remaining 96.7% pointed to third-party domains: media, forums, institutions, platforms and competitors.

YOUR BRAND IS NOT IN THE CONVERSATION. IT IS BEING MENTIONED BY OTHERS, OR IT IS NOT THERE.

BRAND SITES DOMINATE CITATIONS

PikaSEO, compiling studies published through March 2026, holds that brands' own sites are the dominant source on every platform, at between 38% and 70% depending on the model.²¹

Semrush reports that half of ChatGPT's links point directly to business or service sites.²⁰

NO BRAND DOMAIN EXCEEDED 10% IN ITS CATEGORY

In our study, the set of a brand's domains gathers barely 3.3% of citations on average, with a maximum of 9.5%, against 96.7% for third parties.

If someone hears that company sites are 50% of citations and their dashboard shows 3%, the natural, and wrong, conclusion is that their team is failing.

DASHCRAB HYPOTHESIS 02

Both numbers are correct. They measure different things, and the confusion is misaligning expectations across the whole industry.

"Brand sites" is simply the sum of every commercial site in the world: a company, its competitors, and any company that happens to be relevant to the question. Our own taxonomy confirms it from the other side: between 76.4% and 92.3% of each model's citations go to company sites. That half of ChatGPT's citations go to corporate sites says nothing about how many go to yours.

The question that matters to whoever pays is a different one: of the citations in my category, how many are mine? And there the observed ceiling stayed below 10% in every case.

The academic literature backs this reading: the work by Chen, Wang, Chen and Koudas documents a systematic and strong bias in AI search engines toward authoritative third-party sources, over brands' own content and social, in contrast with Google's more balanced mix.¹ The same work points to a structural bias toward large brands that niche players must climb over. CapstonAI, over 24,800 answers, reaches a compatible conclusion: brands' own domains rarely make the top 100 most cited sources.¹⁸

The practical consequence is not to give up. It is to stop treating your own site as the main channel for AI visibility and start treating it as one among several, alongside presence in media, on video platforms and in the institutional sources of the category.

518 SEARCHES, FOUR SURFACES, THE SAME SCREEN

The question of whether SEO predicts citation has answers ranging from 6.8% to 97%. This experiment explains why, and the explanation is more boring and more useful than we expected.

WHAT SHARE OF AI CITATIONS COMES FROM THE ORGANIC TOP 10?

Studies published in 2025 and 2026 • same question, irreconcilable answers



Sources: seoClarity and SEO.com¹²; Ahrefs¹¹; BrightEdge⁹; ALM Corp²³; Moz¹⁰. Different methodologies, surfaces and windows. They are not directly comparable, and that is precisely the point.

EXPERIMENT DESIGN

Between July and August 2026 we ran 518 searches on Google, covering about ten industries and two markets: Argentina and the United States. From each search we captured four surfaces of the same results screen and the same session: the AI Overview, the organic results, the paid ads, and AI Mode after asking a follow-up. In 452 cases (87.3%) there was at least one follow-up question. Total: 10,913 citations.

WHY THE DESIGN MATTERS

The discrepancy in the previous chart is born, to a large extent, from comparing AI citations captured at one moment against positions measured at another, on different searches and different surfaces. Here the overlap is computed per page, within the same search. If a page was in the organic results of that screen and appears cited in the AI Overview of that same screen, it counts. If not, it does not. That quickly removes most of the methodological noise.

47.45%

OF AI OVERVIEW CITATIONS WERE ALREADY IN THE ORGANIC RESULTS OF THAT SAME SEARCH

19.98%

OF ORGANIC PAGES SURVIVE AS A CITATION WHEN MOVING INTO AI MODE

0%

OF PAID ADS MADE IT TO AI MODE, ACROSS THE TEN INDUSTRIES

THE AI OVERVIEW REPEATS YOUR SEO. AI MODE DISCARDS IT.

FROM THE RESULTS SCREEN TO AI MODE	PAGES PRESENT	SURVIVE AS A CITATION	%
AI Overview citations	3,058	794	25.96%
Organic results	3,224	644	19.98%
Paid ads	75	0	0%

If you were in the organic results, the citation survives one time in five. If you were already cited in the AI Overview, one in four. And AI Mode does not just discard: it also multiplies. The average goes from 7.53 citations per AI Overview to 14.98 per AI Mode turn, and 85.29% of those citations were neither in the organic results, nor in the ads, nor in the AI Overview of that same search.

The most plausible reason is that a follow-up is almost never the same question. It usually seeks to go deeper or get concrete: which one suits me if I bill under a certain amount, does this work in my country, how do I set it up step by step. Even if it is born in the same session, the problem changed, and the sources that served the generic question do not necessarily serve the specific one.

YOUR SEO GETS YOU INTO THE AI OVERVIEW. IT GUARANTEES YOU NOTHING IN AI MODE.

And more than half of AI Overview citations (52.6%) come neither from the organic results nor from the ads.

WHAT KIND OF CONTENT LIVES ON EACH SURFACE

There is an established idea that the generative answer favors forums and encyclopedias. In this sample, that is not what happens.

SOURCE TYPE	AI OVERVIEW	AI MODE	ORGANIC
Social (Instagram, Facebook, LinkedIn, TikTok)	8.36%	5.98%	6.10%
YouTube	7.72%	4.86%	1.72%
Reddit	1.13%	1.21%	2.06%
Wikipedia	0.44%	0.17%	0.56%
Reviews and rankings (G2, Trustpilot)	0.05%	0.10%	0.04%

YouTube plus social add up to 16.08% of AI Overview citations. Reddit and Wikipedia together add up to 1.57%. That is an order-of-magnitude difference, not a nuance. The most revealing detail is in the organic column: YouTube weighs 1.72% there and jumps to 7.72% in the AI Overview. More than four times. The AI Overview does not inherit SEO's proportions: it actively overweights video and social content. That is consistent with the external finding that YouTube mentions are the single signal with the highest correlation with AI visibility, at 0.737.¹⁴

A reasonable reading: a video usually answers the question in a format the model can summarize with confidence, with a transcript, chapters and a demonstrative tone. If your category lends itself to that, a well-structured video can weigh more than an article optimized the old-fashioned way.

DASHCRAB HYPOTHESIS 03

The fourteen-fold range between studies is not error: it is surface and category. And both can be measured.

Our experiment produces both extremes of the debate within the same design: 47.45% for the AI Overview and 19.98% for AI Mode. When a study reports 76% or 97%, it is looking at AI Overviews. When another reports 12% or 17%, it is looking at AI Mode or ChatGPT. They do not contradict each other: they describe different surfaces of the same system, and the label "AI citations" throws them all into the same bag.

The second half of the explanation is category, and there the variation is also enormous.

INDUSTRY	AIO CITATIONS ALREADY IN ORGANIC	ORGANIC THAT REACHES AI MODE
Comparisons (best X)	58.52%	16.05%
Fintech and finance	56.05%	12.50%
Health and medicine	50.21%	24.10%
Marketing, advertising and GEO	49.71%	13.57%
Brand protection	37.20%	21.71%
Fitness and sports	37.01%	2.13%

The sample per industry is limited: read as a trend, not as statistical certainty. Even so, the range is far too wide to attribute to chance.

The pattern has a reasonable reading: in categories with an established body of authority, such as health or regulated finance, there is consensus on which sources are trustworthy, and so ranking well and being cited tend to be the same fight. In niche categories or highly personalized decisions, the model goes looking for the specific source for the case even if that means leaving out what dominates the organic results. Fitness is the extreme: only 2.13% of what ranks survives into AI Mode.

THERE IS NO GLOBAL NUMBER FOR HOW MUCH SEO PREDICTS. THERE IS ONE PER SURFACE AND ANOTHER PER CATEGORY, AND THE DIFFERENCE BETWEEN THEM IS AN ORDER OF MAGNITUDE.

DASHCRAB HYPOTHESIS 04

Ads do not buy entry into the answer. Zero, across ten industries, without a single exception.

Of all the ads captured, barely 0.03% of total AI Overview citations came from a paid one. And when moving into AI Mode, persistence was 0% over 75 paid pages.

The most plausible explanation is about text genre, not platform policy: an ad is written to convert someone who has already decided to click. A citation has to sound like a source that answers the question. They are two different writing jobs, and the model needs the second.

Concrete budget implication: search ad spend is not a lever for AI visibility. If someone justifies ad investment with the argument of presence in generative answers, this data contradicts it directly.

THERE IS NO PERFECT NUMBER. THERE IS A FLOOR TO CROSS.

The industry sells AI optimization checklists as if every extra point brought a citation closer. We crossed those scores against real citations and the relationship is not the one being promised.

WHAT AN AEO SCORE IS

A grade from 0 to 100 for how ready a page is to be cited. It combines technical signals (speed, structure, accessibility for crawlers), format signals (clear headings, self-contained answers, scannable text) and identity signals (authorship, verifiable presence of the organization). Almost every tool in the field, ours included, computes something similar.

The implicit promise is that raising the score raises the probability of being cited. That is true only in one stretch, and the stretch is shorter than what is being sold. We analyzed 1,486 URLs with a known AEO score and crossed them against the real citations they received. If the checklist promise were true, the citation rate should rise in an orderly way with the grade. It does not.

PAGE GRADE	CITATION RATE
A • 85 or more	11.9%
B • 70 to 84	2.4%
C • 55 to 69	17.5%
D • 40 to 54	2.8%
F • below 40	0%

The floor is real and it is sharp: no page below 40 points received a single citation. That threshold works as an entry filter, and whatever does not cross it simply does not take part in the game. Above the floor, on the other hand, the score stops ordering anything. Pages with a C grade are cited 17.5% of the time and pages with a B, better scored, barely 2.4%: seven times less. There is no staircase that climbs with the score. There is one step and then a noisy plateau.

BELOW 40 YOU DO NOT EXIST. ABOVE 40, THE SCORE STOPS EXPLAINING WHO WINS.

MORE STRUCTURE DOES NOT MEAN MORE CITATIONS

The format component is the one that ages worst. When you sort articles by their structure and presentation score, the relationship with real citation runs in the opposite direction from the expected one.

Lowest-scoring 25%		9.1%
Second block		1.2%
Third block		0.6%
Highest-scoring 25%		0%

Articles grouped into four blocks of 25% each. Bars scaled to the highest rate.

The most likely explanation is not that format hurts, but that the score is capturing something else: pages that score 100 on structure are long, exhaustive articles, the kind a brand produces to rank. And what a model extracts is a short, specific fragment. The cited pages in our dataset average 844 words against 2,119 for the uncited ones.

AUTHORITY IS NOT THE SAME AS HYGIENE

HEAVILY CITED DOMAIN	AI CITATIONS	AVG. AEO GRADE
youtube.com	176	50
wikipedia.org	107	66
instagram.com	92	32
reddit.com	91	23

Reddit scores 23 out of 100 on the checklist and is one of the most cited domains in the study. No on-page score will ever explain that, because what explains it is what kind of source it is, not how the page is built.

CONVERGENCE WITH THE LITERATURE • VERY HIGH

External causal evidence points the same way: technical artifacts do not move the needle.

Ahrefs ran a causal study between August 2025 and March 2026 with 1,885 pages that added structured data against 4,000 control pages: null effect on ChatGPT, null on Google AI Mode and slightly negative on AI Overviews.¹³ Cyrus Shepard's synthesis of 54 experiments, patents and cases places the llms.txt file last among 23 factors, with 2.0 out of 10 on evidence.¹⁹

In the opposite direction, what does correlate is brand signals off the site: Ahrefs, over 75,000 brands, measures 0.664 for brand mentions on the web and barely 0.218 for inbound links, a third of the strength. YouTube mentions reach 0.737, the highest single correlation in the study.¹⁴

DASHCRAB HYPOTHESIS 05

AI readiness scores measure compliance, not authority. They are useful for spotting what is broken, not for predicting who wins.

The operational reading is about resource allocation. Taking a site from 20 to 50 points is cheap, fast and unlocks the possibility of being cited. Taking it from 70 to 95 consumes far more time and, according to this data, buys no additional visibility. That effort pays more invested in the signals that do correlate: presence in media, in video and in the institutional sources of the category.

And a practical warning: some sites show scores close to zero because of crawling or rendering failures, not because of genuinely low quality. Reading a low score as low authority may simply be misreading a technical error.

WHAT GOOGLE SAYS,
AND WHAT CHANGED IN MAY

The terms have history. AEO, answer engine optimization, was born as SEO jargon around 2018 and 2019, for voice search. GEO, generative engine optimization, has an academic origin: the paper by Princeton, Georgia Tech, the Allen Institute and IIT Delhi published in November 2023.²² In 2026 they are used almost interchangeably, with no standard taxonomy.

On May 15, 2026 Google published its first dedicated guide to optimizing for generative features in Search, updated on July 10. The guide explicitly names AEO and GEO, and frames them as part of SEO, not as separate disciplines.^{3, 4}

Five tactics named as unnecessary for Google Search: lms.txt files and other special markup for AI, which Google declares it does not use; artificially fragmenting content into blocks; rewriting aimed at machines rather than people; chasing inauthentic mentions; and the idea that structured data is a requirement for appearing in generative answers.³ The guide clarifies that maintaining an lms.txt is valid if another service uses it: it is simply not a Google lever.

What it does recommend: SEO fundamentals, clear technical structure, unique first-hand content, and measurement through the generative AI performance report in Search Console.

This is the point of greatest convergence between our data and the official documentation. The F-grade floor, where not a single citation was recorded, is the empirical version of "fundamentals first". The drop in the format component is the empirical version of "do not rewrite for the machine". And the null causal result for structured data measured by Ahrefs¹³ is the experimental version of "structured data is not a requirement".

THREE INDEPENDENT METHODOLOGIES, PLATFORM DOCUMENTATION, AN EXTERNAL CAUSAL EXPERIMENT AND MEASUREMENT OVER REAL CITATIONS, ARRIVED AT THE SAME PLACE BY DIFFERENT ROADS.

Scope warning: Google's guide applies to Google Search, including AI Overviews and AI Mode. It is not a universal specification for ChatGPT, Claude or Perplexity, which use different indexes and crawling semantics. Given the overlap of barely 0.104 between models, extrapolating from one to all would be exactly the mistake this study documents.

09 • IMPLICATIONS

WHAT TO DO ON MONDAY

Nine decisions that follow directly from the data above. None of them is a trick.

1. **Treat the AEO score as a filter, not as a goal.** Cross the floor of 40 across the whole site, because below that nothing gets cited. And stop there: the effort of taking pages from 70 to 95 is the most expensive item in the plan and the one with the least evidence.
2. **Measure where you appear, not just whether you appear.** If your dashboard tells you how many times you were cited but not in which place, you are not measuring the asset: you are measuring the noise around the asset.
3. **Do not report AI visibility as a single figure.** The AI Overview and AI Mode give 47.45% and 19.98% overlap with organic in the same experiment, and category moves the number between 37% and 58.5%.
4. **Stop promising one URL for every LLM.** With 3.5% overlap at the exact-page level, that promise cannot be kept. What is sellable and true is: one domain, a portfolio of pages, and measurement model by model.
5. **Treat your product page as citation material, not sales material.** It is between 22.4% and 31.3% of what each model cites, and ChatGPT's main course. Today most of those pages are written to convert someone who has already decided; to be cited they have to answer first.
6. **Separate the Google dashboard from everyone else's.** Google AI Mode consumes 12.1% social and video; Claude, 0.5%. That is a different content playbook, not a percentage tweak.
7. **If your category lends itself to video, treat it as a citation format.** YouTube weighs 1.72% in organic and 7.72% in the AI Overview: more than four times. It is the only format in the study with that asymmetry in its favor.
8. **Take search ads out of the AI visibility justification.** 0.03% of AI Overview citations and 0% persistence into AI Mode, across ten industries.
9. **Stop treating your site as the main channel.** No brand domain exceeded 10% of the citations in its own category. Visibility is built in the third-party sources the model already consults: media, video and the institutional references of the field.

THREE THINGS WE WOULD LIKE TO CLAIM AND STILL CANNOT

We do not demonstrate causality in any direction. Everything above is observational. The literature we cite shares this limitation and its authors say so: a high correlation between brand mentions and citations does not prove that adding mentions causes citations, because it is just as possible that already strong brands earn both.¹⁹

We do not measure revenue. We know which page types win which places; we do not know which citations produce business.

We do not separate by question intent. It is the most obvious confounder we have left, and it is the next pending experiment.

GLOSSARY

THE TERMS IN THIS STUDY, IN PLAIN ENGLISH

TERM	WHAT IT MEANS HERE
Citation	Every time a model shows a link as a source for its answer. If an answer lists seven sources, that is seven citations.
Citation position	The place that source occupies in the list the model shows. The first is the most visible and, in several interfaces, the only one you see without expanding the rest.
URL depth	How many levels you go down from the homepage. site.com is 0; site.com/blog is 1; site.com/blog/article is 2.
Internal page	A page at depth 2 or more: a specific answer, not the homepage or a section.
Company blog	Editorial article published on a company's domain. It is the most cited category in the study and the one with the worst average position.
Product page	Page describing a specific product or service. ChatGPT's main course and the category the previous classification could not see.
AI Overview	The generated summary Google shows above the results the moment you run a search.
AI Mode	Google's follow-up chat, which you move into when you ask again within the same session. It cites almost twice as many sources per turn, and mostly different ones.
AEO score	Grade from 0 to 100 for how ready a page is to be cited. It combines technical, format and identity signals.
Overlap (Jaccard)	How much two lists coincide: what they have in common divided by the total number of distinct elements. It runs from 0 (nothing in common) to 1 (identical).
Long tail	That volume is not concentrated in a few domains: it takes about 500 to reach half of the total.
GEO / AEO	Industry names for the work of appearing in generative answers. Google considers them part of SEO, not separate disciplines.

EXTERNAL SOURCES CITED

Every figure that does not come from our own data is referenced here. Third-party figures are reproduced as their authors published them.

- Chen, M., Wang, X., Chen, K. & Koudas, N. (2025). *Generative Engine Optimization: How to Dominate AI Search*. University of Toronto. arXiv:2509.08919. <https://arxiv.org/abs/2509.08919>
- Chen, M., Wang, X., Chen, K. & Koudas, N. (2026). *Navigating the Shift: A Comparative Analysis of Web Search and Generative AI Response Generation*. EDBT/ICDT 2026, Tampere. arXiv:2601.16858. <https://arxiv.org/abs/2601.16858>
- Google Search Central (2026). *Optimizing your website for generative AI features on Google Search*. Published May 15, 2026; updated July 10, 2026. <https://developers.google.com/search/docs/fundamentals/ai-optimization-guide>
- Southern, M. G. (2026). *Google's New AI Search Guide Calls AEO And GEO Still SEO*. Search Engine Journal, May 15, 2026. <https://www.searchenginejournal.com/googles-new-ai-search-guide-calls-aeo-and-geo-still-seo/575026/>
- Writesonic (2026). *Do AI Engines Cite the Same Sources? A Study of 161,286 Prompts*. May and June 2026 sample. <https://writesonic.com/blog/ai-citation-source-overlap-study>
- Profound. *Answer Engine Citation Overlap Strategy*. Analysis over 100,000 queries. <https://www.tryprofound.com/blog/citation-overlap-strategy>
- Averi (2026), via AuthorityTech. *Analysis of 680 million AI citations; 11% domain overlap between ChatGPT and Perplexity*. March 2026. <https://authoritytech.io/curated/ai-citation-11-percent-platform-overlap-per-engine-audit-2026>
- ZipTie.dev (2026). *How Different AI Platforms Cite the Same Source Differently*. March 2026. <https://ziptie.dev/blog/>
- BrightEdge (2026). *Generative Parser: overlap of AI Overviews citations with the organic top 10*. Data from Feb 2025 to Feb 2026, via SEO Sherpa and Omnibound. <https://seosherpa.com/ai-search-statistics/>
- Moz (2026). *Analysis of 40,000 searches in Google AI Mode; close to 12% of citations matching the organic top 10*. Via SEO Sherpa. <https://seosherpa.com/ai-search-statistics/>
- Ahrefs (2025), via The Digital Bloom. *2026 AI Citation Position & Revenue Report*. Includes 76.1% of AI Overviews URLs in the top 10, 6.82% ChatGPT overlap and 28.3% of most cited pages with no organic visibility. <https://thedigitalbloom.com/learn/ai-citation-position-revenue-report-2026/>
- seoClarity and SEO.com, via ALM Corp (2025). *The Complete Semrush AI Overviews Study*. Reports 97% of citations coming from pages already ranking in the top 20. <https://almcorp.com/blog/semrush-ai-overviews-study-2026-complete-analysis/>
- Ahrefs (2026), causal study on structured data, via Haide Digital. *LLM Ranking Factors: What the AI Citation Data Actually Shows*. 1,885 treated pages against 4,000 controls, Aug 2025 to Mar 2026. <https://haide.digital/resources/llm-ranking-factors>
- Ahrefs (2025 and 2026), study over 75,000 brands, via Machine Relations Research. *AI Search Citation Factors*. Correlations: brand mentions 0.664; inbound links 0.218; YouTube mentions 0.737. <https://machinerelations.ai/research/ai-search-citation-factors-2026>
- Surfer SEO, via SQ Magazine (2026). *AI Overviews Statistics 2026*. Analysis of 36 million AI Overviews and 46 million citations. <https://sqmagazine.co.uk/ai-overviews-statistics/>
- Semrush (2026). *AI Visibility Index*. Via Machine Relations Research. <https://machinerelations.ai/research/ai-search-citation-factors-2026>
- Peec AI, analysis of 30 million sources, via Contently and Search Engine Land (2026). *Top Sources LLMs Cite Most in 2026*. <https://contently.com/2026/04/29/top-sources-llms-cite/>
- CapstonAI (2026). *LLM Citation Benchmark 2026*. 24,800 answers analyzed across four models, first quarter of 2026. <https://capston.ai/benchmark-llm-citations/>
- Shepard, C. / Zyppy Signal (2026), via Digital Applied. *AI Citation Ranking Factors*. Synthesis of 54 experiments, patents and cases; 23 factors weighted by evidence. Published May 7, 2026. <https://www.digitalapplied.com/blog/ai-search-citation-ranking-factors-2026-data-study>
- Semrush (2025). *How Google's AI Mode Compares to Traditional Search and Other LLMs*. <https://www.semrush.com/blog/ai-mode-comparison-study/>
- PikaSEO (2026). *AI Citation Sources by Platform: Complete Breakdown*. Compilation of studies published through March 2026. <https://pikaseo.com/articles/ai-citation-sources-by-platform>
- Aggarwal, P. et al. (2023). *GEO: Generative Engine Optimization*. Princeton University, Georgia Tech, Allen Institute for AI and IIT Delhi. arXiv:2311.09735. <https://arxiv.org/abs/2311.09735>
- ALM Corp (2026). *Google AI Overview Citations From Top-10 Pages Dropped From 76% to 38%*. March 2026. <https://almcorp.com/blog/google-ai-overview-citations-drop-top-ranking-pages-2026/>
- Pew Research Center (2025), via Omnibound. *Only 1% of users click a link inside an AI Overview*. July 2025. <https://www.omnibound.ai/blog/ai-seo-statistics>
- Everything-PR (2026). *How AI Engines Cite the Web: The Six 2026 Studies That Define the Evidence Base*. June 2026. <https://everything-pr.com/how-ai-engines-cite-the-web-the-six-studies-that-define-the-2026-evidence-base>
- Complementary study of our own.** Bedrossian, R. (2026). *518 questions to Google, 10,913 citations later: this is how AI treats your SEO*. Dashcrab, August 5, 2026. Source of the parallel experiment in section 06. <https://dashcrab.app/en/blog/google-ai-overview-citation-seo-ai-mode-no.html>

How to cite this study: Dashcrab GEO Citation Research (2026). *What AI models actually cite: 21,571 citations, five models and a parallel experiment of 518 searches*. Data cutoff: August 11, 2026.

About the data: citations come from monitoring programs active on the platform. No brands, own domains, individual categories or the number of programs involved are identified; every figure is reported aggregated or anonymized. The third-party domains that are named are public platforms of general reach. The page-type and site-type classification corresponds to taxonomy v4.